HOW THE DISTRIBUTION OF SALIENT OBJECTS IN IMAGES INFLUENCES SALIENT OBJECT DETECTION

B. Schauerte *R. Stiefelhagen*

Institute for Anthropomatics
Karlsruhe Institute of Technology
http://cvhci.anthropomatik.kit.edu/~bschauer/

ABSTRACT

We investigate the spatial distribution of salient objects in images. First, we empirically show that the centroid locations of salient objects correlate strongly with a centered, half-Gaussian model. This is an important insight, because it provides a justification for the integration of such a center bias in salient object detection algorithms. Second, we assess the influence of the center bias on salient object detection. Therefore, we integrate an explicit center bias into Cheng's state-of-the-art salient object detection algorithm. This way, first, we quantify the influence of the Gaussian center bias on salient object detection, second, improve the performance with respect to several established evaluation measures, and, third, derive a state-of-the-art unbiased salient object detection algorithm.

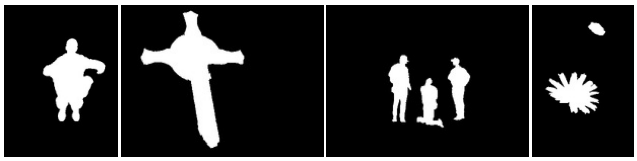
***Index Terms*—** Salient object detection, spatial distribution, photographer bias, center bias

1. INTRODUCTION

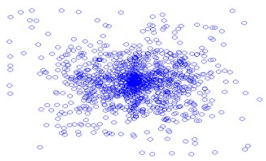
Liu et al. [1] defined the task of *salient object detection* as the binary labeling problem of separating a salient object from the background. A *salient object* is defined as being the object in an image that attracts most of the user's interest. The decision of what is the salient object in an image happens consciously and can be articulated by a person, which stands in contrast to gaze trajectories that occur while a person is viewing an image. Accordingly, since salient objects are likely to attract gaze, salient object detection and gaze prediction are closely related yet substantially different tasks.

The natural tendency of photographers to place the objects of interest near the center of their composition in order to enhance their focus and size relative to the background (see [2]), has been identified as one cause for the often reported center bias in eye-tracking data during eye-gaze studies (see, e.g., [3–5]). As a consequence, the integration of a center bias has become an increasingly important aspect in visual saliency models that focus on gaze prediction (see, e.g., [6–8]). In contrast, most recently proposed salient object detection algorithms do not incorporate an explicit model of such a center bias (see, e.g., [9–12]) and Borji et al. even argued against the integration of such a bias [13][1]. A notable exception and closely related to our work is the work by Jiang et al. [14], in which one of the three main criteria that characterize a salient object is that "it is most probably placed near the center of the image" [14]. The authors justify this characterization with the "rule of thirds", which is one of the most well-known principles of photographic composition (see, e.g., [15]), and use a Gaussian distance metric as a model.

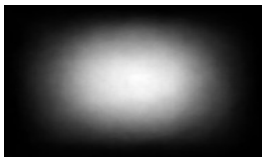
(a) Example images



(b) Segmentation masks of the images shown in Fig. 1(a)



(c) Location scatter plot



(d) Mean segment mask

Fig. 1. Illustration of the Achanta/Liu data set: example images 1(a), the corresponding segmentation masks 1(b), the mean over all segmentation masks 1(d), and the scatter plot of the centroid locations across all images 1(c).

The work presented in this paper was supported by the Quaero Programme, funded by OSEO, French State agency for innovation.

[1] It is important to note that Borji et al. did not consider the effect of potential implicit center biases that are already part of the algorithms.

We go beyond following the rule of third and show that the distribution of the objects' centroids correlates strongly positively with a centered, half-Gaussian distribution. In consequence, we are able to provide a strong empirical justification for integrating a center bias into salient object detection algorithms. To demonstrate the importance, we adapt Cheng's salient object detection algorithm [12] to quantify the influence of the photographer's center bias on salient object detection.

2. MODEL

To investigate the spatial distribution of salient objects in photographs, we use the segmentation masks by Achanta et al. [9, 10] that mark the salient objects in 1000 images of the salient object data set by Liu et al. [1]. More specifically, we use the segmentation masks to determine the centroids of all salient objects in data set and analyze the centroids' spatial distribution. The images in the data set by Liu et al. [1] have been collected from a variety of sources, mostly from image forums and image search engines. 9 users marked the salient objects using (rough) bounding boxes and the salient objects in the image database have been defined based on the "majority agreement".

The Center: Our model is based on a polar coordinate system that has its pole at the image center. Since the images in Achanta's data set have varying widths and heights, we use in the following normalized Cartesian image coordinates in the range $[0,1]\times[0,1]$. The mean salient object centroid location is $[0.5021, 0.5024]^{\mathrm{T}}$ and the corresponding covariance matrix is $[0.0223, -0.0008; -0.0008, 0.0214]$. Thus, we can motivate the use of a polar coordinate system that has its pole at $[0.5, 0.5]^{\mathrm{T}}$ to represent all locations relative to the expected distribution's mode.

The Angles are Distributed Uniformly: Our first model hypothesis is that the centroids' angles in the specified polar coordinate system are uniformly distributed in $[-\pi, \pi]$.

In order to investigate the hypothesis, we use a Quantile-Quantile (Q-Q) plot as a graphical method to compare probability distributions (see [16]). In Q-Q plots the quantiles of the samples of two distributions are plotted against each other. Thus, the more similar the two distributions are, the better the points in the Q-Q plot will approximate the line $f(x) = x$. We calculate the Q-Q plot of the salient object location angles in our polar coordinate system versus uniformly drawn samples in $[-\pi, \pi]$, see Fig. 2 (left). The apparent linearity of the plotted Q-Q points supports the hypothesis that the angles are distributed uniformly.

The Radii follow a Half-Gaussian Distribution: Our second model hypothesis is that the radii of the salient object locations follow a half-Gaussian distribution. We have to consider a truncated distribution in the interval $[0, \infty]$, because the radius – as a length – is by definition positive. If we consider the image borders, we could assume a two-sided truncated distribution, but we have three reasons to work with a one-sided model:

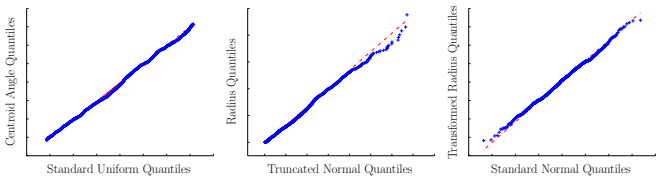


Fig. 2. Quantile-Quantile (Q-Q) plots of the angles versus a uniform distribution (left), radii versus a half-Gaussian distribution (middle), transformed radii (see Sec. 3.1) versus a normal distribution (right).

The variance of the radii seems sufficiently small, the "true" centroid of the salient object may be outside the image borders (i.e., parts of the salient object can be truncated by the image borders), and it facilitates the use of standard statistical tests (see Sec. 3.1).

We use a Q-Q plot against a half-Gaussian distribution to graphically assess the hypothesis, see Fig. 2 (middle). The linearity of the points suggests that the radii are distributed according to a half-Gaussian distribution. The visible outliers in the upper-right are caused by less than 30 centroids that are highly likely to be disturbed by the image borders. Please be aware of the fact that it is not necessary to know the half-Gaussian (or standard Gaussian) distribution's model parameters when working with Q-Q plots (see [16]).

3. EVALUATION

As for the graphical investigation of our hypotheses using Q-Q plots (see Fig. 2), we use the manually annotated segmentation masks by Achanta et al. [9, 10], see Sec. 2, to further investigate our model hypotheses and quantify the influence of the Gaussian center bias on salient object detection.

3.1. Empirical: Hypothesis Analysis

Although often data analysts prefer to use graphical methods such as Q-Q plots to assess the feasibility of a model (see Sec. 2), formal statistical hypothesis tests remain the most important method to disprove hypotheses. The goal of statistical tests is to determine if the (null) hypothesis can be rejected. Consequently, statistical tests either reject (prove false) or fail to reject (fail to prove false) a null hypothesis. But, they can never prove it true (i.e., failing to reject a null hypothesis does not prove it true). However, we can disprove alternate hypotheses and, additionally, we can use a set of statistical tests that are based on different principles. If all tests fails, we have an indicator that the hypothesis is potentially true.

We can quantify the observed linearity in the Q-Q plots (see Fig. 2) to analyze the correlation between the model distribution and the data samples using probability plot correlation coefficients (PPCC) [16]. The PPCC is the correlation coefficient between the paired quantiles and measures the agreement of the fitted distribution with the observed data. The closer the

correlation coefficient is to one, the higher the positive correlation and the more likely the distributions are shifted and/or scaled versions of each other. Furthermore, by comparing against critical values of the PPCC (see [17] and [16]), we can use the PPCC as an additional test, which is closely related to the Shapiro-Wilk test. Furthermore, we can use the correlation to test the hypothesis of no correlation by transforming the correlation to create a t-statistic.

The Angles are Distributed Uniformly: We use Pearson's χ^2 test [18] as a statistical hypothesis test against a uniform distribution. The test fails to reject the hypothesis at significance level $\alpha = 0.05$ ($p = 0.2498$). Considering the circular type of data, we use Rayleigh's and Rao's tests for circular uniformity and both tests fail to reject the hypothesis at $\alpha = 0.05$ ($p = 0.5525$ and $p > 0.5$, respectively; see [19]).

On the other hand, we can – for example – reject the alternative hypotheses of a normal or exponential distribution using the Lilliefors test [20] ($p = 0$ for both distributions[2]).

The obvious linearity of the Q-Q plot, see Fig. 2 (left), is reflected by a PPCC of 0.9988[3], which is substantially higher than the critical value of 0.8880 (see [17]) and thus the hypothesis can not be rejected. Furthermore, the hypothesis of no correlation is rejected at $\alpha = 0.05$ ($p = 0$).

The Radii follow a half-Gaussian Distribution: In order to use standard statistical hypothesis tests, we transform the polar coordinates in such a way that they represent the same point with a combination of positive angles in $[0, \pi]$ and radii in $[-\infty, \infty]$. According to our hypothesis, the distribution of the transformed radii should follow a normal distribution with its mode and mean at 0, see Fig. 2 (right).

Again we disprove exemplary alternate hypotheses: The uniform distribution is rejected by the test against the critical value of the PPCC as well as by Pearson's χ^2 test at $\alpha = 0.05$ ($p = 0$). The exponential distribution is rejected by Lilliefors test at $\alpha = 0.05$ ($p = 0$).

We perform the Jarque-Bera, Lilliefors, Spiegelhalter's, and Shapiro-Wilk test (see [21], [20], [22] and [23]) to test our null hypothesis that the radii have been sampled from a normal distribution (unknown mean and variance). Subsequently, we use a T-test to test our hypothesis that the mean of the radius distribution is 0. The Jarque-Bera, Lilliefors, Spiegelhalter's, and Shapiro-Wilks tests fail to reject the hypothesis at significance level $\alpha = 0.05$ ($p = 0.8746$, $p = 0.2069$, $p = 0.2238$, and $p = 0.1022$, respectively). The correlation that is visible in the Q-Q plot, see Fig. 2 (middle and right), is reflected by a PPCC of 0.9987, which is above the critical value of 0.9984 (see [16]). The hypothesis of no correlation is rejected at $\alpha = 0.05$ ($p = 0$). Furthermore, it is likely that the mode of the (transformed) radius is 0, because the corresponding T-test fails to reject the hypothesis at significance level $\alpha = 0.05$ with $p = 0.9635$.

[2]We report $p = 0$, if the tabulated values are 0 or the Monte Carlo approximation returns 0 or ε (double-precision).

[3]Mean of several runs with $N = 1000$ uniform randomly selected samples.

Fig. 3. An example of the influence of the center bias on segmentation-based salient object detection. Left-to-right: Image, region contrast without and with center bias (RC and RC+CB, resp.), and locally debiased region contrast without and with center bias (LDRC and LDRC+CB, resp.).

3.2. Experimental: Salient Object Detection

Center Biased Saliency Model: We adapt the region contrast model by Cheng et al. [12]. Cheng's model is particularly interesting, because it already provides state-of-the-art performance, which is partially caused by an implicit center bias. Thus, we can observe how the model behaves if we remove the implicit center bias, which was neither motivated nor explained by the authors, and add an explicit Gaussian center bias. We modify the spatially weighted region contrast saliency equation (Eq. 1; see Eq. 7 in [12]) by introducing an explicit center bias term (Eq. 2):

$$S_S(r_k) = w_B \Big[\sum_{r_k \neq r_i} \hat{D}_s(r_k; r_i) w(r_i) D_r(r_k; r_i) \Big] \quad (1)$$

$$+ w_C \, g(C(r_k); \sigma_x, \sigma_y) \quad \text{with} \quad (2)$$

$$\hat{D}_s(r_k; r_i) = \exp(-D_s(r_k; r_i)/\sigma_s^2) \quad . \quad (3)$$

Here, we use a convex combination to control the strength of the influence of the center bias, i.e. $w_B + w_C = 1$ ($w_B, w_C \in \mathbf{R}_0^+$). $\hat{D}_s(r_k; r_i)$ is the spatial distance between regions r_k and r_i, where σ_s controls the spatial weighting. Smaller values of σ_s influence the spatial weighting in such a way that the contrast to regions that are farther away contributes less to the saliency of the current region. The spatial distance between two regions is defined as the Euclidean distance between the centroids of the respective regions using pixel coordinates that are normalized to the range $[0, 1] \times [0, 1]$. Furthermore, $w(r_i)$ is the weight of region r_i and $D_r(\cdot;\cdot)$ is the color distance metric between the two regions (see [12] for more details). Here, the number of pixels in r_i is used as $w(r_i) = |r_i|$ to emphasize color contrast to bigger regions. $C(r_k)$ denotes the centroid of region r_k and g is defined as follows

$$g(x, y; \sigma_x, \sigma_y) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left\{ -\frac{1}{2} \frac{x^2}{\sigma_x^2} \right\} * \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left\{ -\frac{1}{2} \frac{y^2}{\sigma_y^2} \right\} . \quad (4)$$

Interestingly, the unnormalized Gaussian weighted Euclidean distance used by Cheng et al. [12] causes an implicit Gaussian-like center bias, see Fig. 4, because it favors regions whose distances to the other neighbors are smaller. Unfortunately, this has not been motivated, discussed, or evaluated by Cheng et al. In order to remove this implicit bias, we introduce a normalized, i.e. locally debiased, distance function $\check{D}_s(r_k; r_i)$

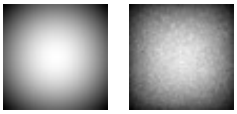


Fig. 4. Illustration of the implicit center bias in the method by Cheng et al. [12]. Left: Each pixel shows the distance weight sum, i.e. $\sum_{r_i} \hat{D}_s(r_k;r_i)$, to all other pixels in a regular grid. Right: The average weight sum depending on the centroid location calculated on the Achanta/Liu data set.

that still weights close-by regions higher than further away regions, but does not lead to an implicit center bias

$$\check{D}_s(r_k;r_i) = \frac{\hat{D}_s(r_k;r_i)}{\sum_{r_i} \hat{D}_s(r_k;r_i)}, \quad (5)$$

$$\text{i.e.} \quad \forall r_k : \sum_{r_i} \check{D}_s(r_k;r_i) = 1 \quad . \quad (6)$$

Baseline Algorithms: In order to compare our results, we use a set of state-of-the-art salient object detection algorithms: The Frequency-Tuned model (FT) by Achanta et al. [9], the Bonn Information-Theoretic Saliency model (BITS) by Klein et al. [11], the Maximum Symmetric Surround Saliency (MSSS) model by Achanta et al. [10], and the Region Contrast (RC) model by Cheng et al. [12] that uses Felzenszwalb's image segmentation method [24]. The latter is the original algorithm we adapted in Sec. 3.2. In order to investigate the influence of the implicit center bias in the region contrast model, we calculate the performance of the locally debiased region contrast (LDRC) model without and with explicit center bias (LDRC and LDRC+CB, respectively). For comparison, we also evaluate the region contrast model with the additional explicit center bias (RC+CB). As additional baseline, we provide the results for simple segment-based and pixel-based – i.e., using Eq. 4 for each pixel with respect to the image center distance and constant variance – center bias models, i.e. $w_C = 1$ (CB_S and CB_P, respectively).

Measures: We can use the binary segmentation masks for saliency evaluation by treating the saliency maps as binary classifiers. At a specific threshold t we regard all pixels that have a saliency value above the thresholds as positives and all pixels with values below the thresholds as negatives. By sweeping over all thresholds $\min(S) \leq t \leq \max(S)$, we can evaluate the performance using common binary classifier evaluation measures.

Most commonly, precision-recall curves are used – e.g., by Achanta et al. [9, 10], Cheng et al. [12], and Klein et al. [11] – to evaluate the salient object detection performance. We use five evaluation measures to quantify the performance of the algorithms. We calculate the area under curve (AUC) of the (interpolated) precision-recall curve (PR) and the receiver operating characteristic (ROC) curve [25]. Complementary to the PR AUC, we calculate the maximum F_1 and $F_{\sqrt{0.3}}$ scores. F_β with $\beta = \sqrt{0.3}$ has been proposed by Achanta et al. to weight precision more than recall for salient object detection [9]. Additionally, we calculate the hit-rate (HR) that measures how often the pixel with the maximum saliency belongs to the salient object.

Method	F_1	F_β	$\int$PR	$\int$ROC	HR
LDRC+CB	0.8034	0.8183	0.8800	0.9624	0.9240
RC+CB	0.7973	0.8120	0.8833	0.9620	0.9340
RC	0.7855	0.7993	0.8710	0.9568	0.9140
LDRC	0.7574	0.7675	0.8302	0.9430	0.8680
BITS	0.7342	0.7582	0.7589	0.9316	0.7540
MSSS	0.7165	0.7337	0.7849	0.9270	0.8420
FT	0.5995	0.6009	0.6261	0.8392	0.7100
CB_S	0.5793	0.5764	0.5920	0.8623	0.6980
CB_P	0.5604	0.5452	0.5638	0.8673	0.7120

Table 1. The maximum F_1 score, maximum F_β score, PR AUC ($\int$PR), ROC AUC ($\int$ROC), and Hit-Rate (HR) of the evaluated algorithms (sorted ascending by F_β).

Results: The performance of RC drops substantially if we remove the implicit center bias as is done by LDRC, see Tab. 1. However, if we add our explicit center bias model to the unbiased model, the performance is substantially increased with respect to all evaluation measures. Furthermore, with the exception of HR, the performance of LDRC+CB and RC+CB is nearly identical with a slight advantage for LDRC+CB. This indicates that we did not lose important information by debiasing the distance metric (LDRC+CB vs RC+CB) and that the explicit Gaussian center bias model is advantageous compared to the implicit weight bias (LDRC+CB and RC+CB vs RC). Most interestingly, LDRC is the best model without center bias, which makes it interesting for applications in which the image data can not be expected to have a photographer's center bias (e.g., image data of surveillance cameras or autonomous robots).

4. CONCLUSION

We investigated the spatial distribution of salient objects using a combination of graphical methods and statistical tests. As a result, we have shown strong empirical evidence that the spatial distribution follows a centered half-Gaussian model. This is an important insight, because it provides an empirical justification for the bias towards the image center that can be found in many salient object detection algorithms. To further investigate the influence, we explicitly integrated the center bias model in a state-of-the-art salient object detection algorithm. This way, we could show that the center bias has a significant, positive influence on the performance and were able to, first, improve the state-of-the-art in salient object detection and, second, derive an unbiased salient object detection algorithm.

5. REFERENCES

[1] T. Liu, J. Sun, et al., "Learning to detect a salient object," in *Proc. CVPR*, 2007.

[2] P.-H. Tseng, R. Carmi, I. G. M. Cameron, D. P. Munoz, and L. Itti, "Quantifying center bias of observers in free viewing of dynamic natural scenes," *Journal of Vision*, vol. 9, no. 7, 2009.

[3] P. Reinagel and A. M. Zador, "Natural scene statistics at the centre of gaze," in *Network: Computation in Neural Systems*, 1999, pp. 341–350.

[4] D. Parkhurst and E. Niebur, "Scene content selected by active vision," *Spatial Vision*, vol. 16, no. 2, pp. 125–154, 2003.

[5] B. W. Tatler, "The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions," *Journal of Vision*, vol. 7, no. 14, 2007.

[6] Y. Yang, M. Song, N. Li, J. Bu, and C. Chen, "What is the chance of happening: a new way to predict where people look," in *Proc. ECCV*, 2010.

[7] T. Judd, K. Ehinger, F. Durand, and A. Torralba, "Learning to predict where humans look," in *Proc. ICCV*, 2009.

[8] A. Borji, D. N. Sihite, and L. Itti, "Probabilistic learning of task-specific visual attention," in *Proc. CVPR*, 2012.

[9] R. Achanta, S. Hemami, F. Estrada, and S. Süsstrunk, "Frequency-tuned Salient Region Detection," in *Proc. CVPR*, 2009.

[10] R. Achanta and S. Süsstrunk, "Saliency detection using maximum symmetric surround," in *Proc. ICIP*, 2010.

[11] D. A. Klein and S. Frintrop, "Center-surround divergence of feature statistics for salient object detection," in *Proc. ICCV*, 2011.

[12] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu, "Global contrast based salient region detection," in *Proc. CVPR*, 2011.

[13] A. Borji, D. N. Sihite, and L. Itti, "Salient object detection: A benchmark," in *Proc. ECCV*, 2012.

[14] H. Jiang, J. Wang, Z. Yuan, T. Liu, and N. Zheng, "Automatic salient object segmentation based on context and shape prior," in *Proc. BMVC*, 2011.

[15] Y. Luo and X. Tang, "Photo and video quality evaluation: Focusing on the subject," in *Proc. ECCV*, 2008.

[16] NIST/SEMATECH, *Engineering Statistics Handbook*, 2012.

[17] R. M. Vogel and C. N. Kroll, "Low-flow frequency analysis using probability-plot correlation coefficients," *Journal of Water Resources Planning and Management*, vol. 115, no. 3, pp. 338–357, 1989.

[18] K. Pearson, "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," *Philosophical Magazine*, vol. 50, no. 302, pp. 157–175, 1900.

[19] E. Batschelet, *Circular statistics in biology*, vol. 24, 1981.

[20] H. W. Lilliefors, "On the kolmogorov-smirnov test for normality with mean and variance unknown," *Journal of the American Statistical Association*, vol. 62, no. 318, pp. pp. 399–402, 1967.

[21] A. K. Bera and C. M. Jarque, "Efficient tests for normality, homoscedasticity and serial independence of regression residuals," *Economics Letters*, vol. 5, no. 3, pp. 255–259, 1980.

[22] D. J. Spiegelhalter, "Diagnostic tests of distributional shape," *Biometrika*, vol. 70, no. 2, pp. 401–409, 1983.

[23] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, pp. 591–611, 1965.

[24] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Comput. Vision*, vol. 59, pp. 167–181, 2004.

[25] J. Davis and M. Goadrich, "The relationship between precision-recall and roc curves," in *Proc. Int. Conf. Machine Learning*, 2006.